(U) Distributed Sensor Fusion Performance Analysis Under an Uncertain Environment

October 2012

Yuwei Liao, Jiangying Zhou, Karen Zachery

Teledyne Scientific Company
5001 S. Miami Blvd, Suite 200, Durham, N.C. 27703

ABSTRACT

Distributed multi-sensor fusion has been widely used in military and civilian applications. In the statistical sensor fusion domain, the design of an optimal fusion processor usually requires the joint statistics of the local sensor outputs. When accurate joint statistical knowledge is not readily available, popular solutions are either to estimate the joint statistics from training data or to simply assume independence of the data. Although it is well known that a fusion solution constructed using empirical data or simplified assumptions often cannot reach the optimal performance, little research has been focused on analyzing the performance difference. This paper presents a systematic analysis of distributed sensor fusion performance in an uncertain operating environment using a Bayesian likelihood ratio fusion model. For the problem where joint statistics of the local sensor outputs cannot be obtained accurately, the sub-optimal fusion processor is assumed to have an estimated correlation coefficient and its performance difference from the optimal scenario is derived analytically using a Gaussian model. We use the detectability index, which fully characterizes the receiver operating characteristic (ROC) curve for the Gaussian model, as the performance metric to compare the optimal and sub-optimal cases. The ratio of detectability indices for the sub-optimal and optimal cases is derived as a function of the true correlation coefficient, the estimated value, and the performance difference between individual local sensors. We prove that the closer the individual local sensor performances, the less vulnerable the fusion performance is to a mismatched estimation of the correlation coefficient. Furthermore, we show that for the special case where all local sensors have the same performance, the optimal fusion performance is always achieved regardless of the estimation deviation from the true correlation coefficient. We provide discussions on the deeper physical meaning of such phenomena. For non-Gaussian sensor noise models, we extend our analysis via computer simulation and provide experimental validations using application specific sensor data of military relevance (e.g., multispectral, hyperspectral). Our results show that similar conclusions hold for a family of heavy-tailed non-Gaussian distribution models.

Keywords: Distributed sensor fusion; Signal detection; Likelihood ratio test; Fusion performance analysis; Robustness

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
OCT 2012	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Distributed Sensor Fusion Performance Analysis Under an Uncertain Environment	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Teledyne Scientific Company 5001 S. Miami Blvd, Suite 200, Durham, N.C. 27703**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM202976. 2012 Joint Meeting of the Military Sensing Symposia (MSS) held in Washington, DC on October 22-25, 2012.

14. ABSTRACT

Distributed multi-sensor fusion has been widely used in military and civilian applications. In the statistical sensor fusion domain, the design of an optimal fusion processor usually requires the joint statistics of the local sensor outputs. When accurate joint statistical knowledge is not readily available, popular solutions are either to estimate the joint statistics from training data or to simply assume independence of the data. Although it is well known that a fusion solution constructed using empirical data or simplified assumptions often cannot reach the optimal performance, little research has been focused on analyzing the performance difference. This paper presents a systematic analysis of distributed sensor fusion performance in an uncertain operating environment using a Bayesian likelihood ratio fusion model. For the problem where joint statistics of the local sensor outputs cannot be obtained accurately, the sub-optimal fusion processor is assumed to have an estimated correlation coefficient and its performance difference from the optimal scenario is derived analytically using a Gaussian model. We use the detectability index, which fully characterizes the receiver operating characteristic (ROC) curve for the Gaussian model, as the performance metric to compare the optimal and suboptimal cases. The ratio of detectability indices for the sub-optimal and optimal cases is derived as a function of the true correlation coefficient, the estimated value, and the performance difference between individual local sensors. We prove that the closer the individual local sensor performances, the less vulnerable the fusion performance is to a mismatched estimation of the correlation coefficient. Furthermore, we show that for the special case where all local sensors have the same performance, the optimal fusion performance is always achieved regardless of the estimation deviation from the true correlation coefficient. We provide discussions on the deeper physical meaning of such phenomena. For non-Gaussian sensor noise models, we extend our analysis via computer simulation and provide experimental validations using application specific sensor data of military relevance (e.g., multispectral, hyperspectral). Our results show that similar conclusions hold for a family of heavy-tailed non-Gaussian distribution models.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**17**	

1.0 Introduction

Sensor fusion, the study of optimal information processing in distributed multi-sensor environments through intelligent integration of multi-sensor data, has gained popularity in recent years due to the increasing demand for more accurate information, more practical and robust procedures to manage data efficiently, and improved system reliability and performance [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. Data fusion systems are now used extensively for both military and civilian applications such as automated target recognition, remote sensing, battle-field surveillance, automated threat recognition systems, manufacturing processes monitoring, robotics and medical applications. Numerous algorithms have been developed and have achieved significant improvement in signal detection performance compared with that of a single sensor.

In the statistical sensor fusion domain, the design of an optimal fusion processor usually requires the joint statistics of the local sensor outputs. When accurate joint statistical knowledge is not readily available, popular solutions are either to estimate the joint statistics from training data or to simply assume independence of the local sensor outputs. Although it is known that a fusion solution constructed using empirical data or simplified assumptions often cannot reach the optimal performance, little research has been focused on analyzing the performance difference. Some published literatures evaluated sensor fusion performance under an uncertain environment where the fusion processor did not have joint statistics. But these literatures mostly used application data or benchmark data sets [11, 12, 13]. There is published research comparing the performance of different algorithms including the naive Bayes, decision trees, and support vector machine (SVM), using several different performance metrics such as accuracy and area under the receiver operating characteristic (ROC) [14]. However, these work lack theoretical analysis that can provide insight into the reasons for the differences in performance of the various algorithms.

In this paper, we present a systematic analysis of distributed sensor fusion performance in an uncertain operating environment, using a Bayesian likelihood ratio fusion model. When the joint statistics of the local sensor outputs cannot be obtained accurately, we define the sub-optimal fusion processor as the one that has an estimated correlation coefficient. The naive Bayesian model is a special case of this type of processor assuming zero correlation coefficient. We use the detectability index, which fully characterizes the receiver operating characteristic (ROC) curve for the Gaussian model, as the performance metric to compare the optimal and sub-optimal cases. We derive the performance difference between the sub-optimal and optimal scenarios analytically using a Gaussian model. The difference is defined as the ratio of detectability indices for the sub-optimal and optimal cases and is derived as a function of the true correlation coefficient, the estimated value, and the performance difference between individual local sensors. We also extend our analysis via computer simulation to non-Gaussian sensor noise models and provide experimental validations using application specific sensor data of military relevance (e.g., multispectral, hyperspectral).

2.0 Preliminaries

The problem we study here is a binary hypothesis testing problem assuming N local sensors, $\boldsymbol{Z}i$, i = 1, 2, ..., N. Each sensor receives observation data and uses its own detection algorithm to generate local likelihood values. These local outputs $\boldsymbol{r}_i$'s are sent to a fusion processor for better detection performance. The Bayesian optimal fusion processor uses all the local outputs to form the likelihood ratio:

$$\lambda = \frac{P(r_1, r_2, ...r_N, | H_1)}{P(r_1, r_2, ...r_N, | H_0)}, \qquad \begin{matrix} H_1 : \text{Signal present} \\ H_0 : \text{Signal absent} \end{matrix} \tag{1}$$

The fusion processor can reach the optimal fusion performance only if the joint statistics between the local outputs under each hypothesis is known. For analysis purpose, we define three types of fusion processor: One is the Optimal Fusion Processor which has the full knowledge of the joint statistics of the local data. The second is the Mismatched Correlation Fusion Processor, which uses estimated correlation

coefficients (mismatched from the true values) to implement the fusion process. The third one, the Independent Fusion Processor is a special case of the Mismatched Correlation Fusion processor, which always assumes the local data is independent. The input data vector to the fusion processor is defined as:

$$\mathbf{r} = [r_1, r_2, \ldots, r_N]^T \tag{2}$$

$$E[\mathbf{r} \mid H_1] = \mathbf{m}_1 = [E[r_1 \mid H_1],\, E[r_2 \mid H_1], \ldots, E[r_N \mid H_1]]^T \tag{3}$$

$$E[\mathbf{r} \mid H_0] = \mathbf{m}_0 = [E[r_1 \mid H_0],\, E[r_2 \mid H_0], \ldots, E[r_N \mid H_{01}]]^T \tag{4}$$

We also assume that the covariance matrices under the two hypotheses are equal, i.e.,

$$\mathbf{\Lambda} = Cov(\mathbf{r} \mid H_1) = Cov(\mathbf{r} \mid H_0) \quad = E[(\mathbf{r} - \mathbf{m}_1)(\mathbf{r}^T - \mathbf{m}_1^T) \mid H_1) \tag{5}$$

3.0 Gaussian Model

We first assume that the local likelihood values are jointly normal. The joint probability density function of the Gaussian vector ***r*** under each hypothesis H_i , is:

$$P(r_1, r_2, ..., r_N \mid H_i) = [(\frac{1}{2\pi})^{\frac{N}{2}} \mid \mathbf{\Lambda} \mid^{\frac{1}{2}}]^{-1} e^{[-\frac{1}{2}(\mathbf{r}^T - \mathbf{m}_i^T)\mathbf{\Lambda}^{-1}(\mathbf{r} - \mathbf{m}_i)]} \tag{6}$$

By definition, the output of the optimal Bayesian log-likelihood ratio fusion processor is:

$$\eta = \ln(\lambda) = \ln \frac{P(r_1, r_2, ..., r_N \mid H_1)}{P(r_1, r_2, ..., r_N \mid H_0)} = \mathbf{r}^T \mathbf{A}(\mathbf{m}_1 - \mathbf{m}_0) = (\mathbf{m}_1 - \mathbf{m}_0)^T \mathbf{A}\mathbf{r} \tag{7}$$

where $\mathbf{A} = \mathbf{\Lambda}^{-1}$

3.1. Performance Analysis

The optimal Bayesian fusion processor requires the joint statistics of the local sensor data. In many applications where the correlation information is not available, an estimation of the correlation coefficient is used for the implementation of the fusion processor, which we have defined as the Mismatch Correlation Fusion Processor. Next we will analyze the performance difference between the Optimal Fusion Processor and Mismatch Correlation Fusion Processor for the Gaussian model.

3.1.1 Two local sensors

We start our analysis with two local sensor case, i.e., $N = 2$. In this case, the output of the Optimal Fusion Processor is:

$$\eta = \ln \frac{P(r_1, r_2) \mid H_1)}{P(r_1, r_2 \mid H_0)} = \frac{(1 - \rho\sqrt{\frac{d_2}{d_1}})r_1 + (1 - \rho\sqrt{\frac{d_1}{d_2}})r_2}{1 - \rho^2} \tag{8}$$

And the output from the Mismatch Correlation Fusion Processor is:

$$\zeta = \frac{(1 - \rho'\sqrt{\frac{d_2}{d_1}})r_1 + (1 - \rho'\sqrt{\frac{d_1}{d_2}})r_2}{1 - \rho'^2} \tag{9}$$

where ρ' is the estimated correlation coefficient. For the special case that the local sensors are assumed to be independent, the output from the fusion processor (Independent Fusion Processor) is:

$$\psi = r_1 + r_2 \tag{10}$$

Since the local outputs r_i's are jointly normal, the linear combination of these variables is marginally normal. Thus η, ζ and ψ are all normally distributed. We use the ROC curves of the fusion processors to evaluate and compare the fusion performances. Since the output of the fusion processors are normally distributed, their ROCs are completely characterized by a single parameter - detectability index [15]. In the following discussions, we will use this parameter as the performance evaluation index. It has been proved in [16] that the detectability index of the Optimal Fusion Processors is:

$$D_{opt} = \frac{d_1 + d_2 - 2\rho\sqrt{d_1 d_2}}{1-\rho^2} \tag{11}$$

Where d_1 and d_2 are the detectability indices of the local sensors. The detectability index of the Independent Fusion Processor is:

$$D_{ind} = \frac{(d_1 + d_2)^2}{d_1 + d_2 + 2\rho\sqrt{d_1 d_2}} \tag{12}$$

The ratio of performance between the Independent Fusion Processor and Optimal Fusion Processor is:

$$\frac{D_{ind}}{D_{opt}} = \frac{(a+1)^2(1-\rho^2)}{(a+1)^2 - 4\rho^2 a} \tag{13}$$

where a is the ratio of performance between the two local sensors, $a = d_1/d_2$. Without loss of generalities, we assume $a \geq 1$.

1) When the two local sensors have the same performance, i.e., $d_1 = d_2$, $D_{ind}=D_{opt}$. In this case, the Independent Fusion Processor has the same performance as that of the Optimal Fusion Processor, even though it neglects the correlation information between the local outputs. This is an interesting result that we will discuss in more details later.

2) For the case that the two local sensors have different performance, i.e., $d_1 \neq d_2$, the performance ratio varies as a and ρ change. Sine D_{opt} is a monotonically decreasing function of ρ when $\rho \in [0, 1]$, we proved that the range of ρ is $[0, \sqrt{\frac{d_2}{d_1}}] = [0, \sqrt{\frac{1}{a}}]$. It has also been proved in [16] that D_{ind}/D_{opt} reaches its global minimal value when $a = 3$ and $\rho = \sqrt{\frac{1}{3}}$, and in this case :

$$\left(\frac{D_{ind}}{D_{opt}}\right)_{\min} = \frac{8}{9} \qquad D_{opt} = d_1 \tag{14}$$

Figure 1 shows how D_{ind}/D_{opt} changes as a function of two variables, a and ρ. When a, the local performance ratio, is fixed (for each individual curve in Figure 1), the fusion performance ratio D_{ind}/D_{opt} decreases as ρ increases. This is because the deviation of the estimated correlation coefficient $\rho' = 0$ from the true coefficient ρ increases. When ρ is fixed, the closer the performance between two local sensors, i.e., the smaller the parameter a is, the closer the performance between the Independent Fusion Processor and the Optimal Fusion Processor. When the two local sensors have exactly the same performance, the Independent Fusion Processor can reach the optimal fusion performance.

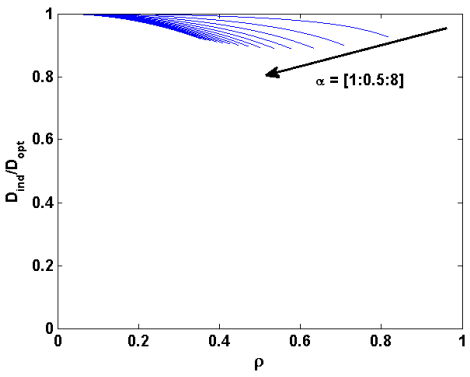


Figure 1

The Independent Fusion Processor is a special case of the Mismatched Correlation Fusion Processor. Next, we will look at the performance difference between the Optimal Fusion Processor and the general Mismatched Correlation Fusion Processor. For the latter, its fusion detectability index is:

$$D_{mismatch} = \frac{\left(d_1 + d_2 - 2\rho'\sqrt{d_1 d_2}\right)^2}{(d_1+d_2)(1+\rho'^2-2\rho\rho') + 2\sqrt{d_1 d_2}(\rho + \rho'^2\rho - 2\rho')} \qquad (15)$$

where ρ is the true correlation coefficient and ρ' is the estimated value that is used to implement the Mismatched Correlation Fusion Processor.

1) For the case that the two local sensors have the same performance, i.e., $d_1 = d_2$,

$$D_{mismatch} = D_{opt} = \frac{2d}{1+\rho} \qquad (16)$$

$D_{mismatch}$ is still independent of the estimated correlation coefficient, and it reaches the same performance as that of the Optimal Fusion Processor.

2) For the case that the two local sensors have different performance, i.e., $d_1 \neq d_2$,

$$\frac{D_{mismatch}}{D_{opt}} = \frac{\left(a+1-2\rho'\sqrt{a}\right)^2\left(1-\rho^2\right)}{\left[(a+1)\left(1+\rho'^2-2\rho\rho'\right)+2\sqrt{a}\left(\rho+\rho'^2\rho-2\rho'\right)\right](1+a-2\rho\sqrt{a})} \qquad (17)$$

where a is stilled defined as the ratio of performance between the two local sensors, $a = d_1/d_2$. The fusion performance ratio $D_{mismatch}/D_{opt}$ is a function of a, ρ and ρ'. Figure 2. - Figure 4 show three examples of $D_{mismatch}/D_{opt}$. In Figure 2, $a = 1.5$, the true correlation coefficient is $\rho = 0.2, 0.4, 0.6$. The estimated correlation coefficient can be in [0, 0.82] (The maximal correlation coefficient that can be is $\sqrt{\frac{d_2}{d_1}} = \sqrt{\frac{1}{a}}$). This example illustrates that when the two local sensors have very close performance, the sub-optimal fusion performance is always very close to the optimal case, regardless how accurate the estimated correlation coefficient is. In Figure 3, the local performance difference increases to $a = 2.0$ and the fusion performance degradation also increases. But the maximal degradation for the independent assumption is always very small as proved earlier in Equation (14), which demonstrates the robustness of the Independent Fusion Processor.

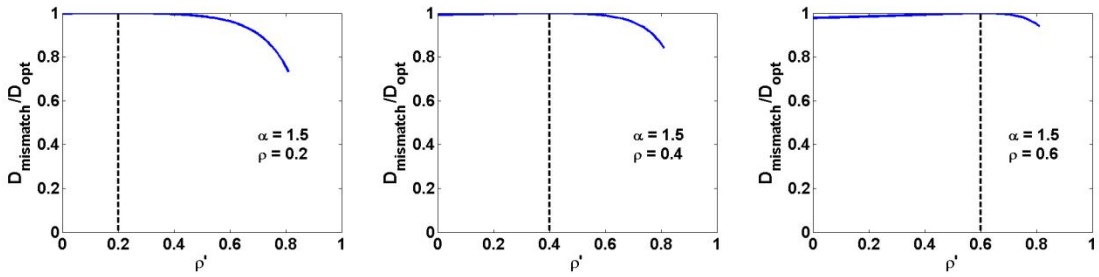


Figure 2. Change of $D_{mismatch}/D_{opt}$ as a function of local sensor performance difference *a*, the true correlation coefficient ρ and the estimated value ρ' - Example 1.

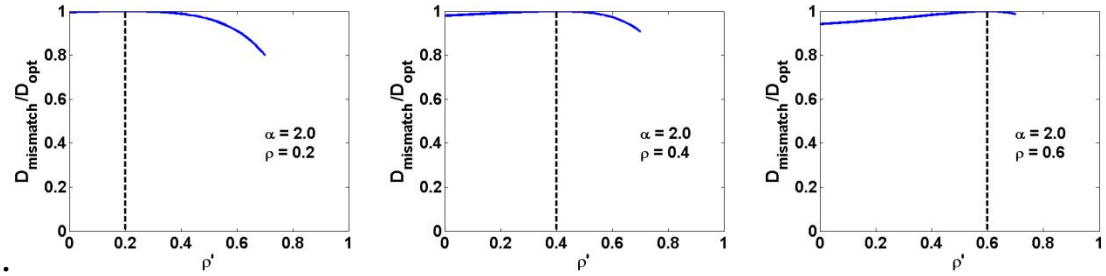


Figure 3. Change of $D_{mismatch}/D_{opt}$ as a function of local sensor performance difference *a*, the true correlation coefficient ρ and the estimated value ρ' - Example 2.

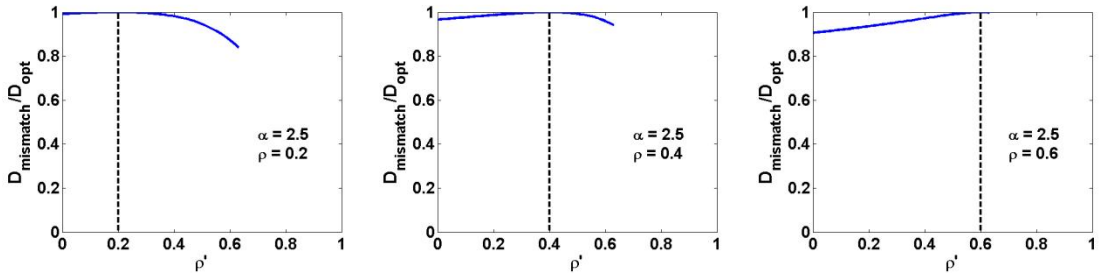


Figure 4 Change of $D_{mismatch}/D_{opt}$ as a function of local sensor performance difference *a*, the true correlation coefficient ρ and the estimated value ρ' - Example 3.

3.1.2 *N* local sensors

We have shown that, for two local sensors with the same performance, the Mismatched Correlation Fusion Processor can reach the optimal performance. The same analysis can be extended to N ($N \geq 3$) local sensor case. Here we assume the output from each sensor is r_i, which is normally distributed under each hypothesis H_1 and H_0. The local outputs are jointly normal, and the correlation coefficient between any two output r_i and r_j ($i \neq j$) is ρ. It is proved in [16] that the Mismatched Correlation Fusion Processor (including the Independent Fusion Processor) has the same fusion performance as that of the optimal one, if all local detectors have the same performance, i.e., $d_1 = d_2 = \ldots = d_N$.

"The same performance" between the optimal and sub-optimal fusion processor refers to the fact that they have the same ROC curve. In the case of Gaussian model, the ROC curve is fully characterized by the detectability index. As for how to reach a specific point on the ROC curve, these two processors require different threshold settings on the fusion output.

3.2. Physical Interpretation

We have proved, using Gaussian model, that the Mismatched Correlation Fusion Processor can reach the same performance as that of the Optimal Fusion Processor when the local sensors have the same

performance. In this section, we will look deeper into the data fusion concept to give a physical interpretation of this result. When the local sensor outputs are statistically independent, the covariance matrix is:

$$\mathbf{\Lambda} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix} \tag{18}$$

And by Equation (7), the output of the Optimal Fusion Processor and its performance is:

$$\eta = \sum_{i=1}^{N} \frac{\Delta m_i r_i}{\sigma_i^2} \qquad D_{opt} = \sum_{i=1}^{N} \frac{(\Delta m_i)^2}{\sigma_i^2} \tag{19}$$

Since $\Delta m_i = \sigma_i^2 = d_i$, we have:

$$\eta = \sum_{i=1}^{N} r_i \qquad D_{opt} = \sum_{i=1}^{N} d_i \tag{20}$$

The physical meaning of the above results is that when the local sensor outputs are independent, the output of the Optimal Fusion Processor is simply the sum of the local outputs and the fusion performance is the sum of the local performance. Data fusion is the fusion of the independent information portion of the local data sources. For the correlated case where the local sensor outputs are not independent, intuitively, the output of the optimal fusion processor is no longer simply the sum of the local outputs and the fusion performance cannot reach the sum of the local performance.

To explain the data fusion in the correlated case, we will use the following results from [15]: If the random variables r_i's are correlated, there always exists a coordinate system in which the random variables are uncorrelated, and the new system is related to the old system by a linear transformation, i.e., there always exists a new set of coordinate axes, $\phi_1, \phi_2, \dots, \phi_N$,

$$\phi_i^T \phi_j = \delta_{ij} \tag{21}$$

such that:

$$r_i' = \mathbf{r}^T \phi_i = \phi_i^T \mathbf{r} \qquad E[(r_i' - m_i')(r_j' - m_j')] = \lambda_i \delta_{ij} \tag{22}$$

where:

$$m_i' = E[r_i'] \qquad Var[(r_i')] = \lambda_i \tag{23}$$

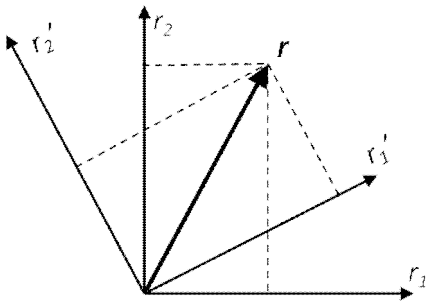


Figure 5. Rotation of the coordinate system. r_1 and r_2 are the coordinates of the original coordinate system, and r_1' and r_2' are the coordinates of the new coordinate system.

In Figure 5., we show the two different coordinate systems. The importance of the new coordinate system is that in this system, the new set of mapped random variables are independent of each other, which means they carry independent information, so that the fusion performance is the sum of the performances of this set of mapped variables. The output of the fusion processor using the new coordinate system is:

$$\eta = \sum_{i=1}^{N} \frac{\Delta m_i' r_i'}{\lambda_i} \qquad D_{opt} = \sum_{i=1}^{N} \frac{(\Delta m_i')^2}{\lambda_i} \tag{24}$$

where:

$$\Delta m_i' = \phi_i^T (\mathbf{m}_1 - \mathbf{m}_0) \tag{25}$$

Next we will use the above results to explain why when the local detectors have the same performance, the Mismatched Correlation Fusion Processor can reach the optimal fusion performance. Since:

$$\lambda\phi = \mathbf{\Lambda}\phi \tag{26}$$

We can see that ϕ_i's are actually the normalized eigenvectors of the covariance matrix $\mathbf{\Lambda}$. In general, the element values of eigenvectors are functions of the local detectability indices d_i's and the correlation coefficient ρ, which means the fusion performance is decided by the local sensor performance and their correlation. For the two sensors case, when the local sensors have the same performance, i.e., $d_1 = d_2 = d$, the covariance matrix is:

$$\mathbf{\Lambda} = \begin{bmatrix} d & \rho d \\ \rho d & d \end{bmatrix} \tag{27}$$

The eigenvalues and eigenvectors are:

$$\lambda_1 = d(1+\rho) \qquad \lambda_2 = d(1-\rho) \tag{28}$$

$$\phi_1\phi_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{29}$$

When the local performances are equal, the rotation of the axes of the is independent of the true correlation coefficient. The Mismatched Correlation Fusion Processor in our earlier discussion processes data in this new rotated coordinate system, and thus can reaches the optimal fusion performance.

For the more general case, when there are N local detectors, if all of them have the same performance, i.e., $d_1 = d_2 = \ldots = d_N = d$, the covariance matrix is:

$$\mathbf{\Lambda} = \begin{bmatrix} d & \rho d & \ldots & \rho d \\ \rho d & d & \ldots & \rho d \\ \vdots & \vdots & \vdots & \vdots \\ \rho d & \rho d & \ldots & d \end{bmatrix} \tag{30}$$

It has two distinct eigenvalues, $\lambda_1 = d[(N\text{-}1)\rho+1]$ and $\lambda_2 = d(1\text{-}\rho)$, where λ_1 is of multiplicity N-1. The eigenvectors, which correspond to the eigenvalue λ_1 and λ_2 are:

$$[\phi_1\phi_2\ldots\phi_N]=\begin{bmatrix} \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{(N-1)+(N-1)^2}} & \frac{1}{\sqrt{(N-2)+(N-2)^2}} & \cdots & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{(N-1)+(N-1)^2}} & \frac{1}{\sqrt{(N-2)+(N-2)^2}} & \cdots & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{(N-1)+(N-1)^2}} & \frac{1}{\sqrt{(N-2)+(N-2)^2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{(N-1)+(N-1)^2}} & -\frac{N-2}{\sqrt{(N-2)+(N-2)^2}} & \cdots & 0 \\ \frac{1}{\sqrt{N}} & -\frac{N-1}{\sqrt{(N-1)+(N-1)^2}} & 0 & \cdots & 0 \end{bmatrix} \quad (31)$$

So for N local detectors case, when the local performances are equal, the rotation of the coordinate system is also independent of the true correlation coefficient ρ. This explains why the Mismatched Correlation Fusion Processor, although using a mismatched correlation coefficient value for data processing, can still reach the optimal fusion performance.

4.0 Non-Gaussian Models

Beyond the analysis using Gaussian model, we also extended our analysis to the non-Gaussian noise models via computer data simulation. We simulated the sensor output as random variables with varied correlation, and then used Bayesian likelihood ratio test as the fusion scheme to compare the fusion performance between the Optimal Fusion Processor and Mismatched Correlation Fusion Processor.

A more general statistical distribution we studied is Gamma distribution, which is a two-parameter family of continuous probability distributions with a shape parameter k and a scale parameter β. Figure 6 shows some examples of Gamma distributions with different shape and scale parameters.

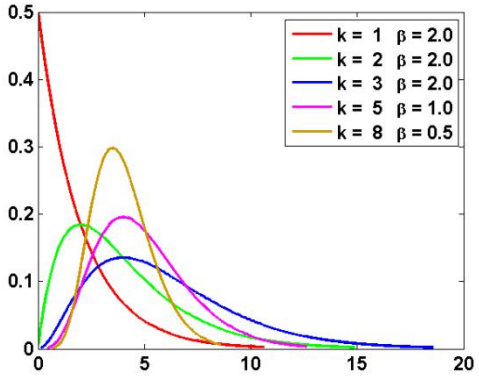


Figure 6. Probability density functions of Gamma distribution with different shape and scale parameters.

We simulated the Gamma distributed random variable X_i's, with certain shape parameter k's, scale parameters β's, and correlation coefficient ρ. We define the inputs to the fusion processor under hypotheses H_1 and H_0 are:

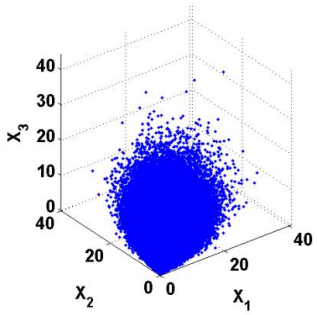
$$r_i=\begin{cases} S_i+X_i & \text{under } H_1 \\ X_i & \text{under } H_0 \end{cases} \quad (32)$$

We implemented two fusion processors using Bayesian likelihood ratio test, one is the Optimal Fusion Processor which we can obtain the joint probability function information of the local outputs using the simulated data, the other is the Independent Fusion Processor which assumes the local outputs are independent. The output from the two fusion processors are:

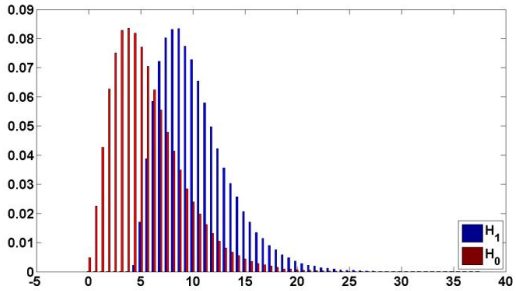
$$\lambda_{opt} = \frac{P(r_1, r_2, \ldots, r_N \mid H_1)}{P(r_1, r_2, \ldots, r_N \mid H_0)} \qquad \lambda_{ind} = \frac{\prod_{i=1}^{N} P(r_i \mid H_1)}{\prod_{i=1}^{N} P(r_i \mid H_0)} \tag{33}$$

The fusion ROC curves are generated to compare the fusion performance difference.

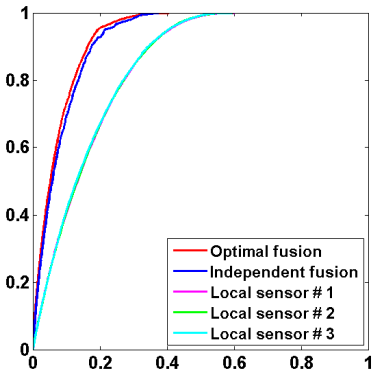
Example 1: Three local sensors (N = 3), each local output has the same Gamma distribution, with k = 3 and β =2 and the same signal to noise ratio. The correlation of the data under each hypothesis is ρ=0.4. Figure 7 (a) shows the simulated Gamma distributed noise data. Figure 7 (b) is the distribution of local outputs under each hypothesis. Figure 7 (c) shows the ROC curves of the individual sensors and the two fusion processors. It proves that when the local sensors have the same performance, the Independent Fusion Processor reaches the optimal fusion processor regardless the non-zero correlation between the data.



(a)



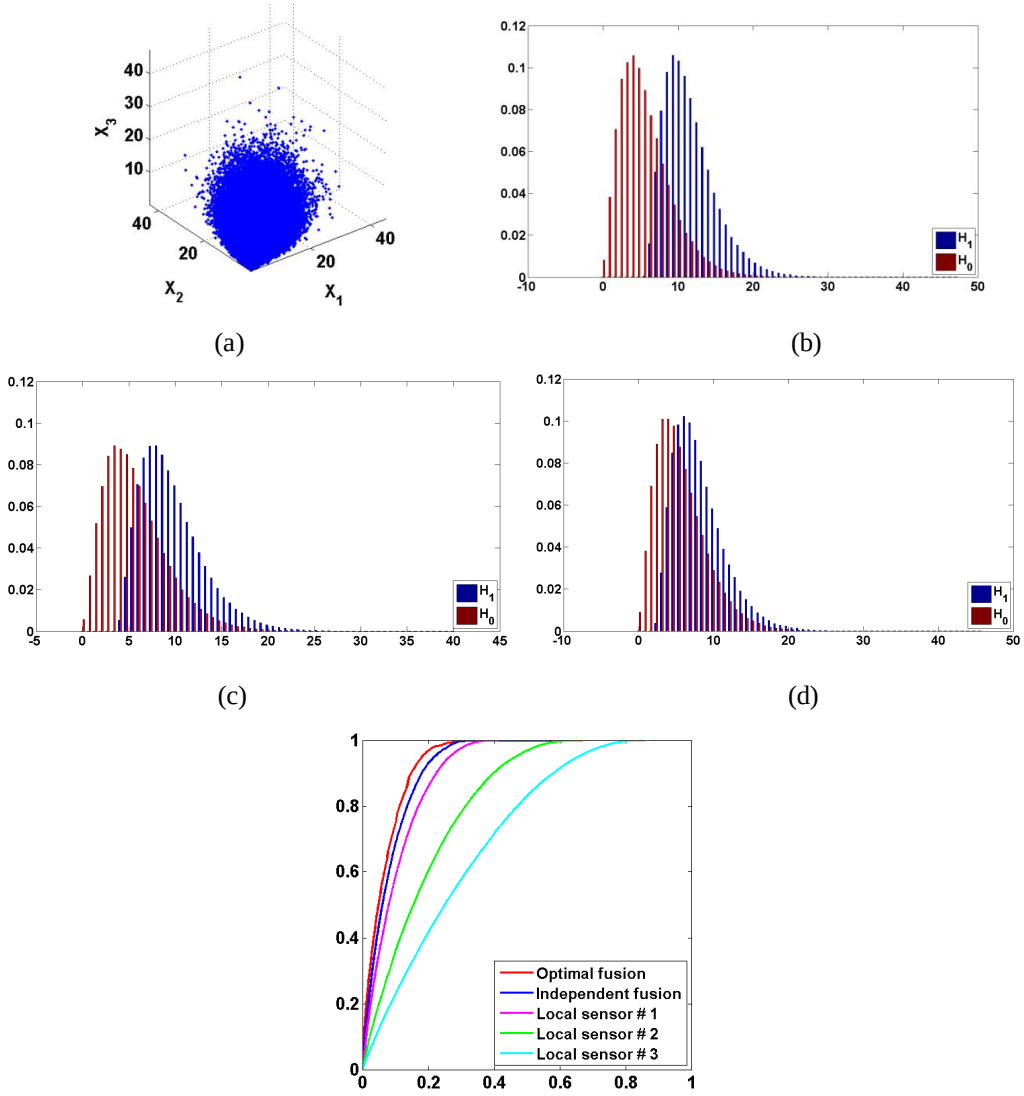
(b)



(c)

Figure 7. Performance comparison between Optimal Fusion Processor and Independent Fusion Processor. All local outputs have the same Gamma distribution with k = 3 and β =2, and the same signal to noise ratio. The correlation of the data under each hypothesis is ρ = 0.4. (a) The simulated Gamma distributed noise data. (b) The distribution of local outputs under each hypothesis. (c) The ROC curves of the individual sensors and the two fusion processors.

Example 2: Three local sensors (N = 3), each local output is Gamma distributed, with k = 3 and β =2, but different signal to noise ratios, thus different local ROC curves. The correlation of the data under each hypothesis is ρ=0.4. Figure 8 (a) shows the simulated Gamma distributed noise data. Figure 8 (b) - (d) are the distributions of local outputs under each hypothesis. Since the signal to noise ratio is different, the local sensors have different detectability indices. Figure 8 (e) shows the ROC curves of the individual sensors and the two fusion processors. It demonstrates that even when the local sensors have different performances, the Independent Fusion Processor still reaches close-to-optimal fusion performance.



(a) (b) (c) (d) (e)

Figure 8. Three local sensors (N = 3), each local output is Gamma distributed, with k = 3 and β =2, but different signal to noise ratios, thus different local ROC curves. The correlation of the data under each hypothesis is ρ=0.4. (a) The simulated Gamma distributed noise data. (b) - (d) : The distributions of local outputs under each hypothesis. (c) The ROC curves of the individual sensors and the two fusion processors.

Example 3: Three local sensors (N = 3), each local output has different distributions. For local sensor #1, it is Gamma distributed with k = 3 and β =2. Local sensor #2 has t-distribution. Local sensor #3 is normally distributed. The correlation of the data under each hypothesis is ρ=0.5. Figure 9 (a) shows the simulated local noise data. Figure 9 (b) - (d) are the distributions of local outputs under each hypothesis.

Figure 9 (e) shows the ROC curves of the individual sensors and the two fusion processors. It shows that even when the local sensors have different distributions, the Independent Fusion Processor still has the robustness in fusing the local sensors.

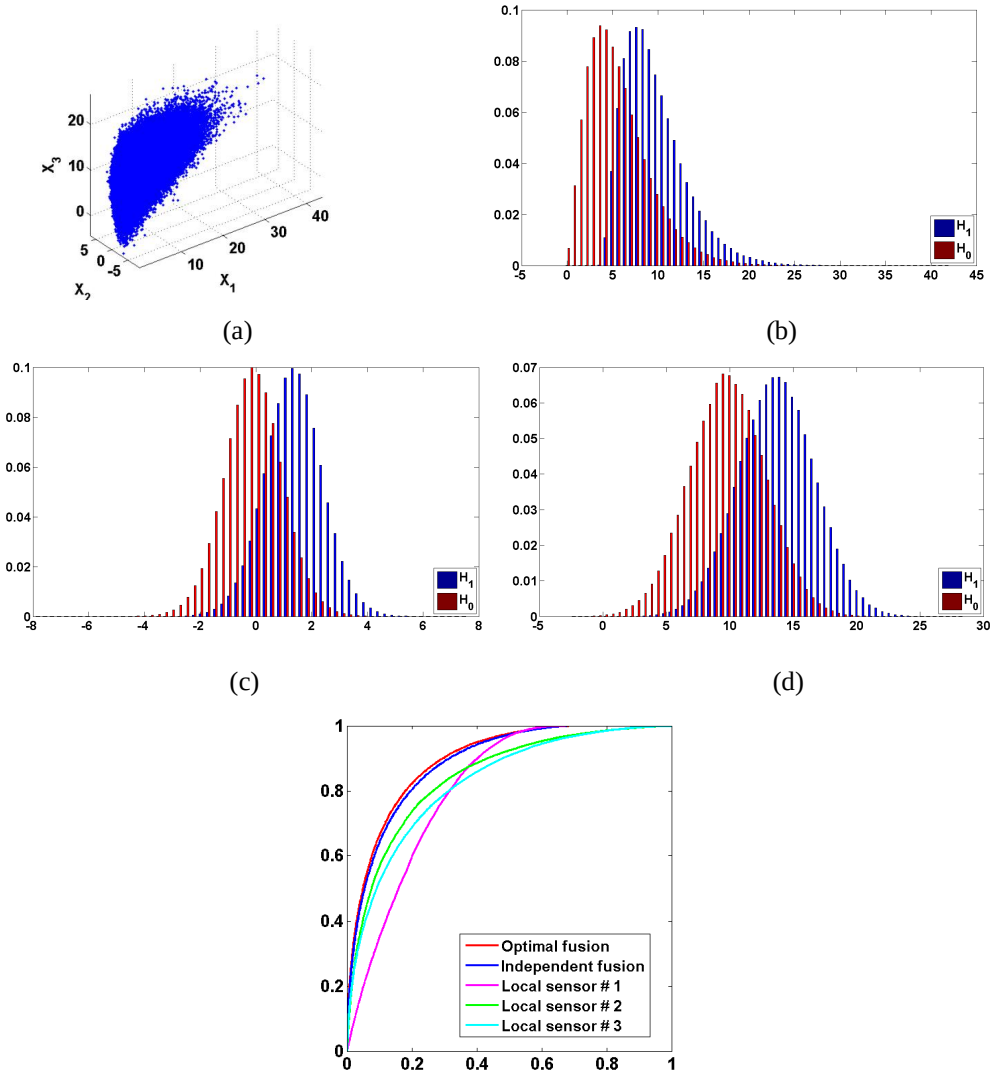


Figure 9. Three local sensors (N = 3), each local output has different distributions. Local sensor #1: Gamma distribution with k = 3 and β =2. Local sensor #2: t-distribution. Local sensor #3: Normal distribution. The correlation of the data under each hypothesis is ρ=0.5. (a) The simulated local noise data. (b) - (d) : The distributions of local outputs under each hypothesis. (c) The ROC curves of the individual sensors and the two fusion processors.

5.0 Experimental data analysis

Since both analytical and simulation results show robustness of the Independent Fusion Processor, we extended our fusion experiments to the data sets collected by the hyperspectral Airborne Signals Intelligence Payload (ASIP) sensor. The ASIP data we analyzed consisted of three passes taken over the same area at different altitudes: low, medium, and high. Our analysis mainly focused on the low altitude data cube which consisted of 130 contiguous spectral bands covering both the VNIR and SWIR regions (400 – 2500 nm). The image size is 334-by-184 pixels per spectral band with a ground resolution of

approximately 0.5 meters. There are five target classes to be detected: metal roof, tarmac, concrete, white plane and blue plane. The sample training pixels for each class is marked as magenta color shown in Figure 10.

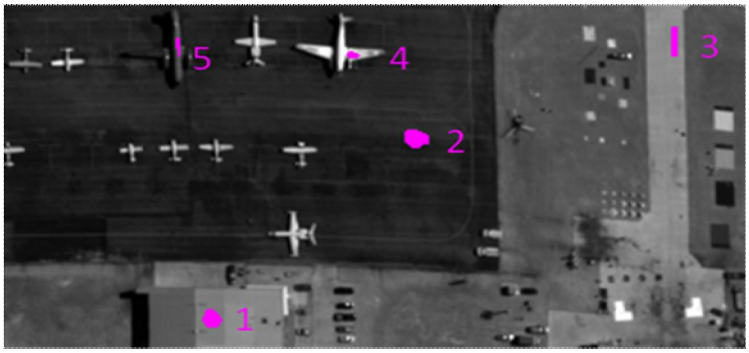


Figure 10. Example frame of one band of hyperspecral data from ASIP sensor. The five classes of interest are: #1: metal roof; #2: tarmac; #3: concrete; #4: white plane; #5: blue plane.

ASIP hyperspectral data was processed through several different classification algorithms to determine the similarity between target and non-target spectra samples. Each classification algorithm was trained with the same spectral library. For each class, the spectral signatures of the sample pixels were averaged to obtain a single representative class signature. These signatures were then used as the spectral library in conjunction with the classification algorithms to match the selected class spectral signatures with similar pixels within the image. There are two main algorithms for detecting each class: the first is to use the features in the spectral domain and then the support vector machine (SVM) as the classification method. The other is to utilize the features in the bispectrum domain and then use different similarity metrics (e.g., Zero-mean normalized cross-correlation, Hausdorff distance, Modified Hausdorff distance) to measure the similarity to the library class signature. Confidence values derived from the classification process along with corresponding ROC curves were then used in our fusion experiments. Given the number of training data available, it is difficult to obtain any meaningful joint probability density function of the local confidence outputs. So, we use Independent Fusion processor to fuse the results from individual detection algorithms for better detection performance. The following two examples show results for two classes, white plane and tarmac.

Example 1: This is an example of detecting white plane class. Algorithms 1 is using Modified Hausdorff distance metrics on bispectral features, and algorithm 2 is using SVM on spectral features. Figure 11 (a) and (b) are the local confidence values of the two algorithms. Figure 11 (c) and (d) show the distribution of local outputs under each hypothesis. Figure 11 (e) and (f) are the joint probability density function of the local outputs assuming we could have all ground truth data for training. Since we have a very small amount of training data, it is not feasible to estimate this joint statistics. The joint statistics are used here only to evaluate the fusion performance of the Optimal Fusion Processor. Figure 11 (g) shows the ROC curves of the individual algorithms and the two fusion processors. We can see that although the joint statistics of the local data is not available, the Independent Fusion Processor still achieved much better detection performance than any of the local algorithms.

Example 2: This is an example of detecting tarmac class. Algorithms 1 is using Hausdorff distance metrics on bispectral features, and algorithm 2 is using modified Hausdorff distance metrics on the same features. Figure 12 (a) and (b) are the local confidence values of the two algorithms. Figure 12 (c) and (d) show the distribution of local outputs under each hypothesis. Figure 12 (e) and (f) are the joint probability density function of the local outputs assuming we could have all data for training. Figure 12 (g) shows the ROC curves of the individual algorithms and the two fusion processors. The tarmac class is more difficult to detect compared with the other classes. So, the performance of the local algorithms are not very good.

But, the Independent Fusion Processor achieved much better detection performance than any local algorithm even though the joint statistics were not readily available.

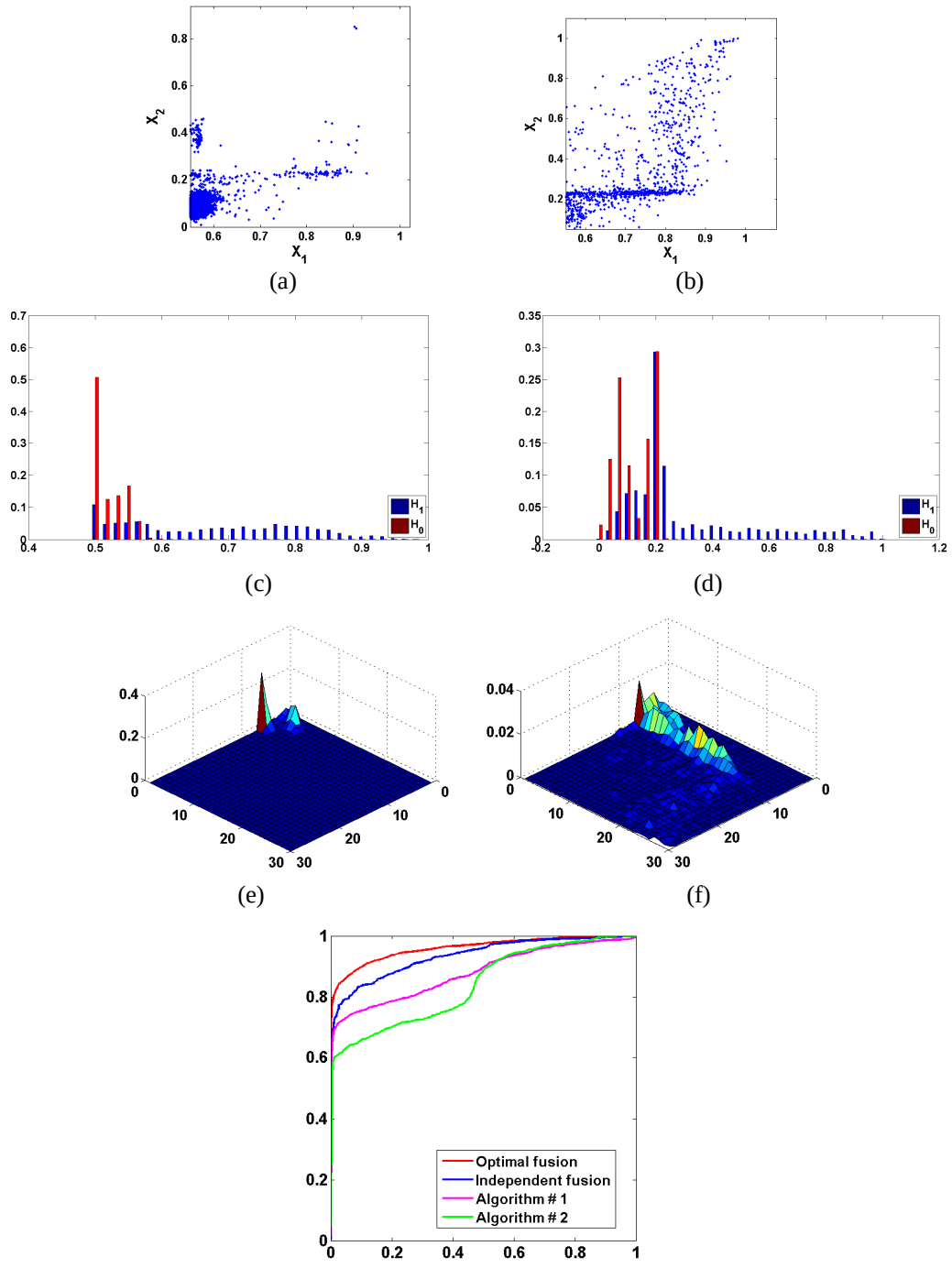


Figure 11. Example of detecting white plane class. Algorithms 1 is using modified Hausdorff distance metrics on bispectral features, and algorithm 2 is using SVM on spectral features. (a) and (b): The local confidence values of the two algorithms. (c) and (d): The distribution of local outputs under each hypothesis. (e) and (f): The joint probability density function of the local outputs under H_0 and H_1. (g): The ROC curves of the individual algorithms and the two fusion processors.

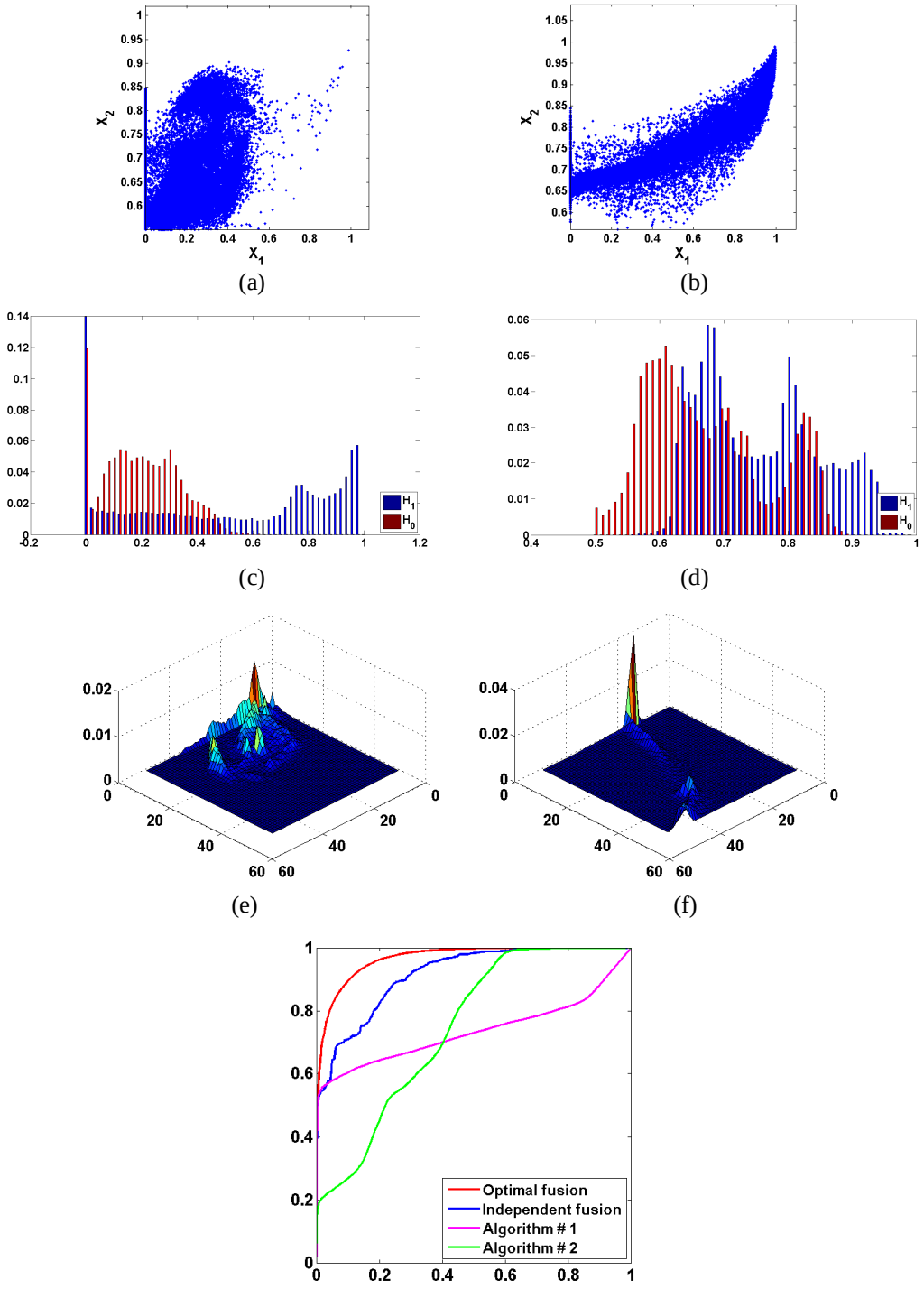


Figure 12. Example of detecting tarmac class. Algorithms 1 is using Hausdorff distance metrics on bispectral features, and algorithm 2 is using modified Hausdorff distance metrics on the same features. (a) and (b): The local confidence values of the two algorithms. (c) and (d): The distribution of local outputs under each hypothesis. (e) and (f): The joint probability density function of the local outputs under H_0 and H_1. (g): The ROC curves of the individual algorithms and the two fusion processors.

6.0 Conclusions

We have conducted a systematic analysis on the distributed sensor fusion performance between the Optimal Fusion Processor and the Mismatch Correlation Fusion Processor which estimates the correlation information of the sensor outputs, using a Bayesian likelihood ratio fusion model. The performance difference is derived analytically for Gaussian distributed data as a function of the true correlation coefficient, the estimated value, and the performance difference between individual local sensors. We proved that when all the local sensors have the same performance, the sub-optimal fusion processor can always reach the optimal fusion performance. A deeper analysis on the physical meaning of such phenomena is also given.

Independent Fusion Processor, which assumes the independence of the local data, is a special case of the Mismatch Correlation Fusion Processor. It can always achieve close-to-optimal fusion performance given the simplicity of implementation. We extended our analysis to non-Gaussian distributed data via computer simulation and provided experimental validations using application specific sensor data. Our results showed that similar conclusions hold for a family of heavy-tailed non-Gaussian distribution models.

In real-world fusion applications, we are often faced with incomplete knowledge and uncertainty regarding the data and sensors. Bayesian interpretation of Occam's razor principle dictates that it is best to use the least complex assumptions when explaining local sensor data. For example, one approach in the choice of fusion schemes is to use fewer degrees-of-freedom to avoid over-fitting the limited available data. Our analytical results provide a theoretical foundation and practical guideline on the trade-off between performance and robustness in the design of a fusion processor.

7.0 References

1. B. V. Dasarathy. *Sensor fusion potential exploitation-innovative architectures and illustrative applications.* Proceedings of the IEEE, 85(1):24–38, Jan. 1997.
2. D. L. Hall and J. Llina. *Handbook of Multisensor Data Fusion*. CRC Press, Boca Raton, FL, 2001.
3. D. L. Hall. *An introduction to multisensor data fusion*. Proceedings of the IEEE, 85(1):6–23, Jan. 1997.
4. Edward Waltz and James Llinas. *Multisensor Data Fusion*. Artech House, Boston, 1990.
5. M. A. Abidi and R. C. Gonzalez. *Data Fusion in Robotics and Machine Intelligence*. Academic Press, Boston, 1992.
6. J. Manyika and H. Durrant-Whyte. *Data Fusion and Sensor Management: a decentralized information-theoretic approach*. Ellis Horwood, New York; London, 1994.
7. L. A. Klein. *Sensor and Data Fusion Concepts and Applications*. SPIE, Bellingham, Wash., 1999.
8. G. D. Riccia, HJ. Lenz, and R. Kruse. *Data Fusion and Perception*. Springer, Wein; NewYork, 2001.
9. G. L. Foresti, C. S. Regazzoni, and P. K. Varshney. *Multisensor Surveillance Systems: The Fusion Perspective*. Kluwer Academic Publishers, Boston, 2003.
10. L. Valet, G. Mauris, and Ph. Bolon. *A statistical overview of recent literature in information fusion.* IEEE Trans. Aerospace and Electronic Systems, AES-16(3):7–14, March 2001.
11. P. Domingos and M. Pazzani. *On the optimality of the simple bayesian classifier under zero-one loss*. Machine Learning, 29:103–130, 1997.
12. I. Rish. *An empirical study of the naive bayes classifier*. Proceedings of IJCAI-01 workshop on Empirical Methods in AI, pages 41–46, 2001.

13. N. B. Amor, S. Benferhat, and Z. Elouedi. *Naive bayes vs decision trees in intrusion detection systems*. Proceedings of the 2004 ACM symposium on Applied computing, pages 420–424, 2004.

14. J. Huang, J. Lu, and C.X. Ling. *Naive bayes vs decision trees in intrusion detection systems*. Proceedings of the Third IEEE International Conference on Data Mining (ICDM'03), pages 553–556, Nov. 2003.

15. H. L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I*. John Wiley & Sons, New York, 1968.

16. Y. Liao, *Distributed Decision Fusion in Signal Detection -- A Robust Approach*. PhD Thesis, Duke University, 2005.